

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 21:40:49 ; Search time 165 Seconds
(without alignments)

35.160 Million cell updates/sec

Title: US-10-633-423-2

Perfect score: 98

Sequence: 1 CFFRGFFNHPYC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	100.0	15	2 AAW23141	AAW23141 NGF recep
2	98	100.0	15	8 ADT89529	Adt89529 Pep5 pept
3	53.5	54.6	453	8 ADN23929	Adn23929 Bacterial
4	53.5	54.6	453	8 ADN23928	Adn23928 Bacterial
5	53.5	54.6	453	8 ADN23930	Adn23930 Bacterial
6	50	51.0	447	8 ADN23797	Adn23797 Bacterial
7	50	51.0	511	8 ADN23700	Adn23700 Bacterial
8	46	46.9	136	6 ABP79789	Abp79789 N. gonorr
9	46	46.9	166	4 AAE09606	Aae09606 Human gen
10	46	46.9	166	5 AAG78978	Ag78978 Human apo
11	46	46.9	166	7 AAE39806	Aae39806 Human gen
12	46	46.9	406	4 AAU30987	Aau30987 Novel hum
13	46	46.9	633	4 AAB61307	Aab61307 Human tra
14	46	46.9	876	4 AAB84959	Aab84959 Apoptin-a
15	46	46.9	876	4 AAB81352	Aab81352 AAP-4 pro
16	46	46.9	1330	3 AAY81947	Aay81947 Heterosig
17	45	45.9	114	4 AAO00257	Aao00257 Human pol
18	44	44.9	415	4 ABB60993	Abb60993 Drosophil
19	44	44.9	1300	4 ABB68075	Abb68075 Drosophil
20	43.5	44.4	96	4 AAB60848	Aab60848 Propionib
21	43.5	44.4	96	6 ABM57367	Abm57367 Propionib
22	43	43.9	60	5 ABP31741	Abp31741 Human ORF
23	43	43.9	67	4 AAO10608	Aao10608 Human pol
24	43	43.9	232	7 ADB64957	Adb64957 Human pro
25	43	43.9	486	7 ADD25149	Add25149 Fertility

ALIGNMENTS

RESULT 1

AAW23141
ID AAW23141 standard; peptide; 15 AA.

XX AC AAW23141;

XX 06-MAR-1998 (first entry)

XX NGF receptor p75-interacting peptide pe2.

XX Protein-protein interaction; interacting polypeptide;

XX polyphage principle; peptide library; nerve growth factor receptor; p75;

XX Circular; cyclic.

XX Synthetic.

XX WO9732017-A1.

XX 04-SEP-1997.

XX 26-FEB-1997; 97WO-EP000931.

XX 26-FEB-1996; 96EP-00102852.

XX (MORP-) MORPHOSYS GES PROTEINOPTIMIERUNG MBH.

XX Ilag V, Ge L;

XX WPI; 1997-448687/41.

XX N-PSDB; AAT79524.

XX Identification of interacting polypeptide encoding nucleic acid sequences
- e.g. to identify protein-protein interactions, which play an important
role in biological processes.

XX Example 4; Page 35; 105pp; English.

XX Cyclic peptides pe2 (AAW23141), pe3 (AAW23142) and pe10 (AAW23143)

XX interact with the p75 intracellular domain of nerve growth factor

XX receptor. They are encoded by oligonucleotides pe2, pe3 and pe10 (see

XX AAT79524-26) that were isolated using a novel method for the

XX identification of nucleic acid sequences encoding 2 or more interacting

XX proteins or peptides. This involves generating 2 libraries of recombinant

XX vectors, expressing members of the libraries in host cells so that at

XX least one interaction is established, and selecting for the generation of

XX a screenable or selectable property representing the interaction of

XX polypeptides. In this case, selection of correct pairing interactions is

XX via selectively inhibiting phage (SIP)

Adn61164 Radish nu
Aaw65118 Human GDN
Abp45110 Human Bly
Adg95937 Single ch
Aaw37462 Human Ret
Adj58711 Human Ret
Aay15178 Human GFR
Aab19583 Human PRO
Aab24051 Human PRO
Aaw65117 Human GDN
Aaw65116 Human GDN
Aaw37463 Human Ret
Aaw84180 A GDNFR-a
Aaw84186 Glial cel
Aay83226 PROS38 Po
Aay84590 Amino aci
Aay15177 Human GFR
Aab19582 Human PRO
Aab24411 Human PRO
Aab00171 PROS38 po

```

XX SQ Sequence 15 AA;
Query Match 100.0%; Score 98; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFNHNPRYC 15
Db 1 CFFRGGFNHNPRYC 15

RESULT 2
ADT89529
XX ADT89529 standard; peptide; 15 AA.
XX AC ADT89529;
XX DT 16-DEC-2004 (first entry)
XX DE Pep5 peptide.
XX KW Nerve regeneration; gene therapy; vaccine; neuroprotective; nootropic;
XX KW Pep5.
XX OS Unidentified.
XX FH Key Location/Qualifiers
XX FT Misc-difference 4 /note= "Encoded by MGN"
XX FT Misc-difference 13 /note= "Encoded by MGN"
XX PN US2004191240-A1.
XX PD 30-SEP-2004.
XX PF 31-JUL-2003; 2003US-00633423.
XX PR 28-MAR-2003; 2003JP-00092923.
XX PR 30-APR-2003; 2003US-00427741.
XX PA (TOHY/) TOHYAMA M.
XX PA (YAMA/) YAMASHITA T.
XX PI Tohyama M, Yamashita T;
XX WPI; 2004-698659/68.
XX DR N-PSDB; ADT89528.
XX CC Regenerating nerves or modulating nerve regeneration comprises inhibiting
XX CC or modulating p75 signal transduction pathway by administering a
XX CC transduction agent, e.g. p21 or Rho, or an agent that interacts with the
XX CC transduction agent.
XX PS Claim 42; SEQ ID NO 2; 209pp; English.
XX CC The present invention relates to a method for regenerating nerves or
XX CC modulating nerve regeneration. The method involves inhibiting or
XX CC modulating a p75 signal transduction pathway. The invention is useful for
XX CC treating preventing or diagnosing neurological diseases based on nerve
XX CC regeneration and for identifying agents useful for nerve regeneration.
XX CC The invention is also useful in gene therapy and for preparing vaccine.
XX CC The present sequence is a Pep5 peptide. This sequence is used in the
XX CC exemplification of the invention.
XX SQ Sequence 15 AA;
Query Match 100.0%; Score 98; DB 8; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.7e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFNHNPRYC 15

```

```

XX DB 1 CFFRGGFNHNPRYC 15
RESULT 3
ADN23929
XX ADN23929 standard; protein; 453 AA.
XX AC ADN23929;
XX DT 02-DEC-2004 (first entry)
XX DE Bacterial polypeptide #6582.
XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.
XX OS Bacteria.
XX PN US2003233675-A1.
XX PD 18-DEC-2003.
XX PF 20-FEB-2003; 2003US-00369493.
XX PR 21-FEB-2002; 2002US-0360039P.
XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.
XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX CC New recombinant DNA construct comprising a promoter positioned to provide
XX CC for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source, useful for producing plants with improved properties.
XX PS Claim 1; SEQ ID NO 6582; 122pp; English.
XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by
XX CC providing improved plant growth and development under at least one stress
XX CC condition, improved lignin production or improved galactomannan
XX CC production. This sequence represents a bacterial polypeptide used in the
XX CC scope of the invention. Note: The sequence data for this patent did not
XX CC form part of the printed specification but was obtained in electronic
XX CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX

```

SQ Sequence 453 AA;

Query Match 54.6%; Score 53.5; DB 8; Length 453;
Best Local Similarity 28.6%; Pred. No. 7.5;
Matches 12; Conservative 3; Mismatches 0; Indels 27; Gaps 2;

QY 1 CFFRGGFN-----HNP-----RYC 15
||:||||:|
Db 160 CFYRGGFYNFEGGSSNFLNPETPGHSPTDGNSTAGRYC 201

RESULT 4
ADN23928

ID ADN23928 standard; protein; 453 AA.

XX AC ADN23928;

DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #6581.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX CC New recombinant DNA construct comprising a promoter positioned to provide
XX CC for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 6581; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with
XX CC improved plant properties, e.g. improved cold, heat or drought tolerance,
XX CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
XX CC increased resistance to plant disease, better growth rate by modification
XX CC of the cell cycle pathway with plant growth regulators, increased rate of
XX CC homologous recombination, modified seed oil or protein yield and/or
XX CC content, improved yield by modification of carbohydrate, nitrogen or
XX CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 453 AA;

Query Match 54.6%; Score 53.5; DB 8; Length 453;
Best Local Similarity 28.6%; Pred. No. 7.5;
Matches 12; Conservative 3; Mismatches 0; Indels 27; Gaps 2;

QY 1 CFFRGGFN-----HNP-----RYC 15
||:||||:|
Db 160 CFYRGGFYNFEGGSSNFLNPETPGHSPTDGNSTAGRYC 201

RESULT 5
ADN23930

ID ADN23930 standard; protein; 453 AA.

XX AC ADN23930;

DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #6583.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX KW pathogen tolerance; pest tolerance; plant disease resistance;
XX KW cell cycle pathway modification; plant growth regulator;
XX KW homologous recombination; seed oil yield; protein yield; carbohydrate;
XX KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX KW bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.

XX PA (HINK/) HINKLE G J.

XX PA (SLAT/) SLATER S C.

XX PA (CHEN/) CHEN X.

XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX DR WPI; 2004-061375/06.

XX CC New recombinant DNA construct comprising a promoter positioned to provide
XX CC for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 6583; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX CC promoter functional in a plant cell, where the promoter is positioned to
XX CC provide for expression of a polynucleotide encoding a polypeptide from a
XX CC microbial source. The invention also relates to a transformed plant
XX CC comprising the recombinant DNA construct and a method of producing a
XX CC transformed plant having an improved property. The plant is a crop plant
XX CC such as maize or soybean. The method of producing a transformed plant
XX CC having an improved property comprises transforming a plant with the
XX CC recombinant DNA construct and growing the transformed plant, where the
XX CC polynucleotide or polypeptide is useful for improving plant properties.
XX CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 453 AA;

Query Match 54.6%; Score 53.5; DB 8; Length 453;
Best Local Similarity 28.6%; Pred. No. 7.5;
Matches 12; Conservative 3; Mismatches 0; Indels 27; Gaps 2;
Qy 1 CFYRGGFPN-----HNP-----RYC 15
Db 160 CFYRGGFPNPEGGSSNFLNPETPGHSPTDGNSCTGAGRYC 201
||:||||:|

RESULT 6

ID ADN23797 standard; protein; 447 AA.

XX AC ADN23797;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #6450.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX FT New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX PS Claim 1; SEQ ID NO 6450; 122pp; English.

XX CC The invention relates to a recombinant DNA construct comprising a
XX promoter functional in a plant cell, where the promoter is positioned to
XX provide for expression of a polynucleotide encoding a polypeptide from a
XX microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition, improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX SQ Sequence 447 AA;

Query Match 51.0%; Score 50; DB 8; Length 447;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFYRGGFPN 9

Db 154 CFYRGGFPN 162
||:||||:

RESULT 7

ID ADN23700

ADN23700 standard; protein; 511 AA.

XX AC ADN23700;

XX DT 02-DEC-2004 (first entry)

XX DE Bacterial polypeptide #6353.

XX KW Recombinant DNA construct; transformed plant; improved plant property;
XX cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
XX pathogen tolerance; pest tolerance; plant disease resistance;
XX cell cycle pathway modification; plant growth regulator;
XX homologous recombination; seed oil yield; protein yield; carbohydrate;
XX nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
XX bacterial polypeptide.

XX OS Bacteria.

XX PN US2003233675-A1.

XX PD 18-DEC-2003.

XX PF 20-FEB-2003; 2003US-00369493.

XX PR 21-FEB-2002; 2002US-0360039P.

XX PA (CAOY/) CAO Y.
XX PA (HINK/) HINKLE G J.
XX PA (SLAT/) SLATER S C.
XX PA (CHEN/) CHEN X.
XX PA (GOLD/) GOLDMAN B S.

XX PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX FT New recombinant DNA construct comprising a promoter positioned to provide
XX for expression of a polynucleotide encoding a polypeptide from a
XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 6353; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

CC promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a

CC microbial source. The invention also relates to a transformed plant

CC comprising the recombinant DNA construct and a method of producing a

CC transformed plant having an improved property. The plant is a crop plant

CC such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the

CC recombinant DNA construct and growing the transformed plant, where the

CC polynucleotide or polypeptide is useful for improving plant properties.

CC The recombinant DNA construct is useful for producing plants with

CC improved plant properties, e.g. improved cold, heat or drought tolerance,

CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,

CC increased resistance to plant disease, better growth rate by modification

CC of the cell cycle pathway with plant growth regulators, increased rate of

CC homologous recombination, modified seed oil or protein yield and/or

CC content, improved yield by modification of carbohydrate, nitrogen or

CC phosphorus use and/or uptake, by modification of photosynthesis or by

CC providing improved plant growth and development under at least one stress

CC condition, improved lignin production or improved galactomanan

CC production. This sequence represents a bacterial polypeptide used in the

CC scope of the invention. Note: The sequence data for this patent did not

CC form part of the printed specification but was obtained in electronic

CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX

SQ Sequence 511 AA;

Query Match 51.0%; Score 50; DB 8; Length 511;

Best Local Similarity 77.8%; Pred. No. 29;

Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFPRGGFFN 9

Db 204 CFVRGGFFN 212

RESULT 8

ABP79789

XX ABP79789 standard; protein; 136 AA.

XX

AC ABP79789;

XX

XX 07-MAR-2003 (first entry)

XX

DE N. gonorrhoeae amino acid sequence SEQ ID 6108.

XX

XX Antibacterial; infection; vaccine; gene therapy.

XX

OS Neisseria gonorrhoeae.

XX

XX WO200279243-A2.

XX

XX 10-OCT-2002.

XX

PP 12-FEB-2002; 2002WO-IB002069.

XX

XX 12-FEB-2001; 2001GB-00003424.

XX

XX (CHIR-) CHIRON SPA.

XX

XX Fontana MR, Piza M, Maignani V, Monaci E;

XX

XX WPI; 2003-058415/05.

DR

DR N-PSDB; ABZ40759.

XX

XX New protein from Neisseria gonorrhoeae, useful for the manufacture of a

PT medicament for treating or preventing N. gonorrhoeae infection.

XX

PS Disclosure; Page 634; 815pp; English.

XX

CC The present invention relates to proteins from Neisseria gonorrhoeae.

CC Also disclosed are the nucleic acid molecules encoding the proteins and

CC antibodies that specifically bind to the proteins. The composition

CC comprising the protein, nucleic acid or antibody is useful for the

CC manufacture of a medicament for treating or preventing N. gonorrhoeae

CC infection, this may be in the form of a vaccine or gene therapy.

CC Sequences given in records ABP76736-ABP81046 represent nucleic acid

CC molecules of the invention

XX

SQ Sequence 136 AA;

Query Match 46.9%; Score 46; DB 6; Length 136;

Best Local Similarity 61.5%; Pred. No. 33;

Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 FRGGFFNHRVC 15

Db 106 FRGGTGRPLC 118

RESULT 9

AAE09606

ID AAE09606 standard; protein; 166 AA.

XX

AC AAE09606;

XX

XX 19-NOV-2001 (first entry)

XX

DE Human gene 14 encoded novel protein HWMEC68, SEQ ID NO:42.

XX

XX Human; cytostatic; gene therapy; inflammatory disorder; neural disorder;

KW Parkinson's disease; Alzheimer's disease; immune system disorder; AIDS;

KW autoimmune disease; rheumatoid arthritis; muscular disorder; ischaemia;

KW reproductive disorder; Crohn's disease; pulmonary disorder; cancer;

KW myocardial infarction; glomerulonephritis; nephrotic syndrome; tumour;

KW haematopoietic disorder; rhinitis; asthma; diabetes; atherosclerosis;

KW drug screening; endocrine disorder; leukaemia.

XX

OS Homo sapiens.

XX

XX

PH Key Location/Qualifiers

FT Misc-difference 157 /label= Unknown

FT /note= "Encoded by MAA"

XX

XX WO200155311-A2.

XX

XX 02-AUG-2001.

XX

XX 17-JAN-2001; 2001WO-US001315.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR

PR 04-FEB-2000; 2000US-0180628P.

PR

PR 24-FEB-2000; 2000US-0184664P.

PR

PR 02-MAR-2000; 2000US-0186350P.

PR

PR 16-MAR-2000; 2000US-0189874P.

PR

PR 17-MAR-2000; 2000US-0190076P.

PR

PR 18-APR-2000; 2000US-0198123P.

PR

PR 19-MAY-2000; 2000US-0205515P.

PR

PR 28-JUN-2000; 2000US-0209467P.

PR

PR 30-JUN-2000; 2000US-0215135P.

PR

PR 07-JUL-2000; 2000US-0216647P.

PR

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR

PR 14-JUL-2000; 2000US-0218290P.

PR

PR 26-JUL-2000; 2000US-0220963P.

PR

PR 26-JUL-2000; 2000US-0220964P.

PR

PR 14-AUG-2000; 2000US-0224518P.

PR

PR 14-AUG-2000; 2000US-0224519P.

PR

PR 14-AUG-2000; 2000US-0225213P.

PR

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0227182P.
PR 22-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0231968P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.

PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-502629/55.
N-PSDB; AAD16540.
New polynucleotides and polypeptides, useful for diagnosing, treating, preventing or prognosing inflammatory, neural, immune system, muscular, reproductive, pulmonary, cardiovascular or proliferative disorders, or cancer.
Claim 11; SEQ ID NO 42; 464pp; English.
AAD16527-AAD16544 represent cDNAs corresponding to novel human protein genes, and AAE09593-AAE09610 represent the proteins they encode. AAD16545 -AAD16572 represent novel human genomic DNA. The novel proteins and their DNAs are useful for diagnosing, treating, preventing and/or prognosing inflammatory disorders (bursitis or tendonitis); neural disorders (e.g. Parkinson's disease, Alzheimer's disease); immune system disorders; AIDS; autoimmune diseases (e.g., rheumatoid arthritis); muscular disorders; reproductive disorders; gastrointestinal disorders (malabsorption syndrome, Crohn's disease); pulmonary disorders; cardiovascular disorders (myocardial infarction, ischaemia, arrhythmias); renal disorders (glomerulonephritis, nephrotic syndrome); cancerous disease and conditions (breast cancer); hyperproliferative disorders (leukemia, hyperplasia); tumours; foetal and developmental abnormalities; haematopoietic disorders; respiratory disorders (rhinitis, asthma); angiogenic disorders; diabetes; atherosclerosis; endocrine disorders; pregnancy-related disorders and infections. The novel protein DNA is useful in gene therapy and drug screening. The proteins can also be used

CC to aid wound healing and epithelial cell proliferation, to prevent skin
 CC aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues, to

Query Match 46.9%; Score 46; DB 4; Length 166;
 Best Local Similarity 40.0%; Pred. No. 41;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNPRYC 15
 Db 114 CYYGDGFFGENNEFC 128

RESULT 10
 ABG78978
 ID ABG78978 standard; protein; 166 AA.

XX AC ABG78978;

XX DT 15-NOV-2002 (first entry)

XX DE Human apoptosis related protein from cDNA clone HMWEC68.

XX KW Human; apoptosis related protein; immunodeficiency;
 KW B cell immunodeficiency; severe combined immunodeficiency; gene therapy;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; asthma;
 KW diabetes mellitus; allergy; inflammatory condition; thrombosis;
 KW graft-versus-host disease; blood-related disorder; atherosclerosis;
 KW hyperproliferative disorder; cancer; renal disorder; arrhythmia;
 KW acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
 KW Goodpasture's syndrome; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; endocrine disorder; Addison's disease;
 KW reproductive system disorder; endometriosis; infection; Crohn's disease;
 KW gastrointestinal disorder.

XX OS Homo sapiens.

XX PN US2002086811-A1.

XX PD 04-JUL-2002.

XX PF 17-JAN-2001; 2001US-00764861.

XX PR 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 28-JUN-2000; 2000US-0214886P.

PR 07-JUL-2000; 2000US-0216647P.

PR 07-JUL-2000; 2000US-0216880P.

PR 11-JUL-2000; 2000US-0217487P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 28-JUL-2000; 2000US-0220984P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 22-AUG-2000; 2000US-0226868P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 05-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 08-SEP-2000; 2000US-0229513P.

PR 21-SEP-2000; 2000US-0231413P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 27-SEP-2000; 2000US-0235834P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.

XX (ROSE/) ROSEN C A.

PA (RUBE/) RUBEN S M.

PA (BARA/) BARASH S C.

XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2002-642253/69.

XX N-PSDB; ABS41130.

XX Novel polypeptide useful for diagnosis, prognosis, prevention, and
 PT treatment of immune, hyperproliferative, renal, respiratory,
 PT cardiovascular, reproductive, endocrine, gastrointestinal and
 PT neurological disorders.

XX Claim 11; Page 197; 258pp; English.

XX The invention relates to 18 apoptosis related proteins (or proteins 90%
 CC similar to them) and their encoding nucleic acids (including fragments,
 CC allelic variants and nucleic acids hybridising to them). Also included
 CC are vectors, host cells, antibodies, method of detecting the activity of
 CC the proteins, and the genes for the proteins. The nucleic acids, proteins
 CC and antibodies are useful for diagnosing and/or treating, prognosing or
 CC preventing immunodeficiencies (e.g. B cell immunodeficiencies, severe
 CC combined immunodeficiencies), autoimmune disorders (rheumatoid arthritis,
 CC multiple sclerosis, diabetes mellitus), allergic reactions and conditions
 CC (e.g. asthma), inflammatory conditions, graft-versus-host disease, blood-
 CC related disorders (thrombosis, atherosclerosis), hyperproliferative
 CC disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis),
 CC cardiovascular disorders (e.g. arrhythmia), respiratory disorders
 CC (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's
 CC disease, Parkinson's disease), endocrine disorders (e.g. Addison's
 CC disease), reproductive system disorders (e.g. endometriosis), infectious
 CC diseases (e.g. viral, bacterial or fungal infections), and
 CC gastrointestinal disorders (e.g. Crohn's disease). The protein is useful
 CC for identifying a binding partner to the polypeptide. The protein is also
 CC useful for stimulating keratinocyte growth, to prevent hair loss, to
 CC modulate mammalian characteristics such as body height, weight, hair
 CC colour, and to increase or decrease storage capabilities, fat content,
 CC lipid, protein, carbohydrate, vitamins, minerals, cofactors or other
 CC nutritional components. The present sequence represents a human apoptosis
 CC related protein of the invention

XX Sequence 166 AA;

Query Match 46.9%; Score 46; DB 5; Length 166;

Best Local Similarity 40.0%; Pred. No. 41;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNPRYC 15

Db 114 CYYGDGFFGENNEFC 128

PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 PR 17-JAN-2001; 2001US-00764861.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM, Barash SC;
 XX
 DR WPI; 2003-765488/72.
 DR N-PSDB; AAD60439.
 XX
 PT New isolated nucleic acids and their encoded polypeptides potentially
 PT useful for diagnosing and treating disease.
 XX
 PS Claim 1; Page 195-196; Opp; English.
 XX
 CC AAD60426-AAD60443 represent cDNAs corresponding to 17 human secreted
 CC protein genes and AAE39793-AAE39810 represent the proteins they encode.
 CC AAD60444-AAD60471 represent human genomic DNAs. The invention is useful
 CC for preventing, treating or ameliorating immunodeficiencies such as
 CC Bruton's disease, Wiskott-Aldrich syndrome and Chediak-Higashi syndrome,
 CC autoimmune disorders such as Hashimoto's thyroiditis, systemic lupus
 CC erythematosus, rheumatoid arthritis, myasthenia gravis, Goodpasture's
 CC syndrome, multiple sclerosis, haemolytic anaemia, Addison's disease and
 CC Sjogren's syndrome, allergic reactions such as asthma, rhinitis and
 CC eczema. The invention is also useful in the treatment of tissue-specific
 CC inflammatory disorders such as adenitis, cholecystitis, keratitis,
 CC retinitis and polymyelitis, CNS disorders such as stroke and ischaemia,
 CC cardio-vascular disorders such as atherosclerosis and myocarditis and
 CC neurodegenerative disorders such as Parkinson's disease and Alzheimer's
 CC disease. The invention is also useful in gene therapy. The present
 CC sequence represents a human secreted protein of the invention
 XX
 SQ Sequence 166 AA;
 Query Match 46.9%; Score 46; DB 7; Length 166;
 Best Local Similarity 40.0%; Pred. No. 41;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CFFRGGFFNNPRYC 15
 |::||| |:
 DB 114 CYYGDGFFGNNNEFC 128
 RESULT 12
 AAU30987
 ID AAU30987 standard; protein; 406 AA.
 XX
 AC AAU30987;
 XX
 DT 18-DEC-2001 (first entry)
 Human transcriptional regulator protein #7.
 Human; transcriptional regulator protein; TXREG.
 Homo sapiens.
 WO200078954-A2.

XX Novel human secreted protein #1478.
 DE
 XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX
 OS Homo sapiens.
 XX
 PW WO200179449-A2.
 XX
 PD 25-OCT-2001.
 XX
 PF 16-APR-2001; 2001WO-US008656.
 XX
 PR 18-APR-2000; 2000US-00552929.
 PR 26-JAN-2001; 2001US-00770160.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2001-611725/70.
 XX
 PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX
 PS Claim 20; Page 386; 765pp; English.
 XX
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU33304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX
 SQ Sequence 406 AA;
 Query Match 46.9%; Score 46; DB 4; Length 406;
 Best Local Similarity 40.0%; Pred. No. 97;
 Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CFFRGGFFNNPRYC 15
 |::||| |:
 DB 317 CYYGDGFFGNNNEFC 331
 RESULT 13
 AAB61307
 ID AAB61307 standard; protein; 633 AA.
 XX
 AC AAB61307;
 XX
 DT 30-MAR-2001 (first entry)
 Human transcriptional regulator protein #7.
 XX
 DE Human; transcriptional regulator protein; TXREG.
 KW Homo sapiens.
 OS
 XX
 PN WO200078954-A2.

XX PD 28-DEC-2000.
XX PF 15-JUN-2000; 2000WO-US016766.
XX PR 18-JUN-1999; 99US-0140109P.
XX PA (TNCY-) INCYTE GENOMICS INC.
XX PI Lal P, Yue H, Tang YT, Baughn MR, Azimzai Y, Tran B;
XX DR WPI; 2001-041425/05.
XX PT Isolated polypeptide with a human transcriptional regulator protein
XX FT sequence is useful for the diagnosis, prevention and treatment of
XX PT disorders associated with the immune, reproductive and cardiovascular
XX PT systems.
XX XX
XX PS Claim 1; Page 98--100; 142pp; English.
XX CC The present invention relates to human transcriptional regulator protein
XX CC (TXREG) sequences. The antagonist and an agonist of the proteins of the
XX CC invention are used to treat disorders associated with decreased or
XX CC increased expression or activity of TXREG
XX CC
XX SQ Sequence 633 AA;

Query Match 46.9%; Score 46; DB 4; Length 633;
Best Local Similarity 40.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNPRYC 15
|::||| |:
Db 53 CYYGDGFFGENNEFC 67

RESULT 14
AAB84959
ID AAB84959 standard; protein; 876 AA.
XX AC AAB84959;
XX XX
XX DT 06-AUG-2001 (first entry)
XX DE Apoptin-associating prtoein (AAP)-4.
XX KW Apoptin-associating protein; AAP; nuclear localization; apoptosis;
XX KW apoptin; p53; cell proliferation; cancer; auto-immune disease; AAP-4;
XX KW cytostatic; immunosuppressive.
XX OS Homo sapiens.
XX PN WO200142461-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-NL000905.
XX PR 10-DEC-1999; 99EP-00204242.
XX PR 07-APR-2000; 2000EP-00250119.
XX PA (LEAD-) LEADD BV.
XX PI Noteborn MHM, Danen-Van Oorschot AAAM, Rohn JL, Weiss B, Toschi L;
XX DR N-PSDB; AAF83859.
XX DR WPI; 2001-367872/38.
XX DR N-PSDB; AAF83859.
XX PT Isolated or recombinant Apoptin-associating proteinaceous (AAP)
XX PT substance, useful for treating cancer and autoimmune disease, localizes
XX PT to the nucleus for induction of apoptosis.
XX XX
XX PS Disclosure; Fig 6; 46pp; English.

XX CC The invention relates to an isolated or recombinant Apoptin-associating
XX CC proteinaceous (AAP) substance capable of providing nuclear localization
XX CC and apoptosis. AAP associates with Apoptin, inducing apoptosis either
XX CC alone or with Apoptin. It is thought to work particularly in transformed
XX CC or tumorous cells. AAP co-localizes with chromatin/DNA structures in the
XX CC nucleus of cells, in an initial apoptotic phase. AAP, nucleic acids
XX CC encoding AAP and vectors and host cells comprising the nucleic acids are
XX CC used to induce apoptosis, especially when it is p53-independent. They are
XX CC used to treat a disease where enhanced cell proliferation or decreased
XX CC cell death is observed, especially cancer or auto-immune disease. The
XX CC present sequence represents the AAP-4 sequence
XX CC
XX SQ Sequence 876 AA;

Query Match 46.9%; Score 46; DB 4; Length 876;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNPRYC 15
|::||| |:
Db 296 CYYGDGFFGENNEFC 310

RESULT 15
AAB83352
ID AAB83352 standard; protein; 876 AA.
XX AC AAB83352;
XX XX
XX DT 09-OCT-2001 (first entry)
XX DE AAP-4 protein sequence.
XX KW Apoptin-associating protein; p53-independent apoptosis induction;
XX KW cell proliferation; cell death; cancer; autoimmune disease; Vector.
XX OS Unidentified.
XX PN EP1118662-A2.
XX XX
XX PD 25-JUL-2001.
XX PF 17-JAN-2001; 2001EP-00200163.
XX PR 17-JAN-2000; 2000EP-00200169.
XX PR 07-APR-2000; 2000EP-00250118.
XX PA (LEAD-) LEADD BV.
XX PI Noteborn MHM, Danen-Van Oorschot AAM, Rohn JL, Weiss B;
XX DR WPI; 2001-477075/52.
XX DR N-PSDB; AAF87098.
XX XX
XX PT Nucleic acids encoding Apoptin-associating proteins, useful for the
XX PT prevention, diagnosis and treatment of cancers and autoimmune disorders.
XX PS Disclosure; Fig 11; 32pp; English.
XX CC This sequence represents the AAP-4 protein sequence. The invention
XX CC relates to nucleic acids encoding Apoptin-associating proteins. The
XX CC nucleic acid, vector containing it, host cell and/or the encoded protein
XX CC may be used to induce p53-independent apoptosis. This may further
XX CC comprise using a nucleic acid encoding Apoptin or functional (or
XX CC functional equivalent/fragment). They may be used in this way for the
XX CC treatment of a disease in which enhanced cell proliferation or decreased
XX CC cell death is observed, especially a cancer or autoimmune disease. The
XX CC nucleic acids may also be used for detecting the presence of cancer cells
XX CC (or cancer-prone cells) in a sample, by transfecting cells in the sample
XX CC with the nucleic acid or vector, and determining the percentage of
XX CC apoptosis in the sample. This may further comprise determining the
XX CC intracellular localisation of a proteinaceous substance derived from the

CC nucleic acid or vector in the cells. The presence of the proteinaceous
CC substance is detected by immuno-staining the cells with the antibody.
CC They may also be used to identify a putative cancer-inducing agent (e.g.
CC a genome or functional fragment) by contacting a cell sample with a
CC candidate agent and detecting the presence of cancer cells (or cancer-
CC prone cells). They may also be used to identify a putative effector of the
CC activity of protein, by binding it with a candidate effector and
CC determining it's binding
XX
SQ Sequence 876 AA;

Query Match 46.9%; Score 46; DB 4; Length 876;
Best Local Similarity 40.0%; Pred. No. 2.1e+02;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 CFFRGGFNNHNPYC 15
|:|:|:|:|:|:|:
Db 296 CYYGDFGFGENNEFC 310

Search completed: October 4, 2005, 22:12:12
Job time : 168 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 22:01:05 ; Search time 39 Seconds
(without alignments)
37.006 Million cell updates/sec

Title: US-10-633-423-2

Perfect score: 98

Sequence: 1 CFFRGGFNFNPRYC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53.5	54.6	453	2 T24127	probable chitinase
2	50	51.0	272	2 T20139	hypothetical prote
3	50	51.0	447	2 T19078	hypothetical prote
4	50	51.0	511	2 T31899	hypothetical prote
5	45	45.9	433	2 H84055	UDP-N-acetylmurama
6	44	44.9	277	2 AD0809	xanthosine phospho
7	44	44.9	365	2 JC7527	nuclear retroviral
8	44	44.9	708	1 A43717	alpha-galactosidas
9	44	44.9	1291	2 T13389	hypothetical prote
10	43	43.9	134	2 S46106	probable membrane
11	43	43.9	373	2 F86382	hypothetical prote
12	43	43.9	660	2 AH2348	hypothetical prote
13	42	42.9	124	2 C96559	F519.3 [imported]
14	42	42.9	439	2 H82875	oligopeptide trans
15	42	42.9	473	2 A56175	adhesive plaque pr
16	42	42.9	1477	2 T18534	protein-tyrosine k
17	41	41.8	237	2 T34473	hypothetical prote
18	41	41.8	373	1 WMBE76	U16 protein - hum
19	41	41.8	501	2 A90595	cardiolipin synthe
20	41	41.8	537	2 H84642	hypothetical prote
21	41	41.8	658	2 T50080	probable DNA repl
22	41	41.8	826	2 G97073	uncharacterized pr
23	40	40.8	143	2 E72349	conserved hypothet
24	40	40.8	273	2 AH2432	hypothetical prote
25	40	40.8	412	2 S62538	hypothetical coile
26	40	40.8	442	2 AB0964	regulatory protei
27	40	40.8	442	2 C41853	hexose phosphate t
28	40	40.8	462	2 T05995	hypothetical prote
29	40	40.8	638	1 KQHUP	plasma kallikrein

30 40 40.8 664 2 C84747 probable protein k
31 40 40.8 713 2 D85503 lysine decarboxyla
32 40 40.8 713 2 D90652 lysine decarboxyla
33 40 40.8 713 2 A20531 lysine decarboxyla
34 40 40.8 713 2 B64743 lysine decarboxyla
35 40 40.8 1339 2 G84764 hypothetical prote
36 39.5 40.3 314 2 B57126 dual specificity p
37 39.5 40.3 609 2 F84824 hypothetical prote
38 39.5 40.3 957 2 E84547 probable disease r
39 39 39.8 97 2 T03940 hypothetical prote
40 39 39.8 117 2 A71014 hypothetical prote
41 39 39.8 125 2 T03283 hypothetical prote
42 39 39.8 131 2 T03279 hypothetical prote
43 39 39.8 132 2 T03282 hypothetical prote
44 39 39.8 154 2 S76191 hypothetical prote
45 39 39.8 155 2 AF2439 tmRNA-binding prot

ALIGNMENTS

RESULT 1

T24127

probable chitinase (EC 3.2.1.14) precursor R10D12.15 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T24127; T25316; T25318

R;Percy, C.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19842

A;Accession: T24127

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-453 <WIL>

A;Cross-references: UNIPROT:P92013; EMBL:Z81109; PIDN:CAB03255.1; GSPDB:GN00023; CESP:R10D12.15

A;Experimental source: clone R10D12

R;Percy, C.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z20014

A;Accession: T25316

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-453 <WI2>

A;Cross-references: EMBL:Z82054; PIDN:CAB04840.1; GSPDB:GN00023; CESP:T26F2.1

A;Experimental source: clone T26F2

A;Accession: T25318

A;Status: preliminary; translated from GB/EMBL/DBDJ

A;Molecule type: DNA

A;Residues: 1-453 <WI3>

A;Cross-references: EMBL:Z82054; PIDN:CAB04842.1; GSPDB:GN00023; CESP:R10D12.15

A;Experimental source: clone T26F2

C;Genetics:

A;Gene: CESP:R10D12.15; CESP:T26F2.1

A;Map position: 5

A;Introns: 23/1; 118/3; 150/2; 237/3; 369/1

C;Keywords: glycosidase; hydrolase

Query Match 54.6%; Score 53.5; DB 2; Length 453;

Best Local Similarity 28.6%; Pred. No. 0.57;

Matches 12; Conservative 3; Mismatches 0; Indels 27; Gaps 2;

QY 1 CFFRGGFNFN-----HNP-----RVC 15

Db 160 CFFRGGFNFNFGGPPSNFLNPETPGHSPDTGNSCTSGAGYC 201

RESULT 2

T20139

hypothetical protein C51E3.8 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T20139

R;Wilkinson, J.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z19227
A:Accession: T20139
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: UNIPROT:O01981; EMBL:Z78410; PIDN:CAB01643.1; GSPDB:GN00023; CESP:CS51E3
C:Genetics:
A:Gene: CESP:CS1E3.8
A:Map position: 5
A:Introns: 38/3; 117/3

Query Match

51.0%; Score 50; DB 2; Length 272;

Best Local Similarity 77.8%; Pred. No. 1.3; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 2;

1 CFFRGGGFN 9

157 CFYRGGFN 165

RESULT 3

T19078

hypothetical protein C08B6.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T19078

R:Wilkinson, J.

submitted to the EMBL Data Library, May 1996

A:Reference number: Z19070

A:Accession: T19078

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-447 <WIL>

A:Cross-references: UNIPROT:Q17816; EMBL:Z72502; PIDN:CAA96587.1; GSPDB:GN00023; CESP:CO08B6

C:Genetics:

A:Gene: CESP:C08B6.4

A:Map position: 5

A:Introns: 17/1; 112/3; 144/2; 363/1

Query Match

51.0%; Score 50; DB 2; Length 447;

Best Local Similarity 77.8%; Pred. No. 2.1; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 2;

1 CFFRGGGFN 9

154 CFYRGGFN 162

RESULT 4

T31899

hypothetical protein T05H4.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T31899

R:Blanchard, M.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid T05H4.

A:Reference number: Z21097

A:Accession: T31899

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-511 <BLA>

A:Cross-references: UNIPROT:O16512; EMBL:AF016452; PIDN:AAB66013.1; GSPDB:GN00023; CESP:CS51E3

C:Genetics:

A:Gene: CESP:T05H4.7

A:Map position: 5

A:Introns: 76/1; 192/2; 229/1; 414/1

Query Match

51.0%; Score 50; DB 2; Length 511;

Best Local Similarity 77.8%; Pred. No. 2.4; Mismatches 0; Indels 0; Gaps 0;

Matches 7; Conservative 2;

1 CFFRGGGFN 9

204 CFYRGGFN 212

RESULT 5

H84055

UDP-N-acetylmuramate-alanine ligase murC [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004

C:Accession: H84055

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: H84055

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-433 <STO>

A:Cross-references: UNIPROT:Q9K7M1; GB:AP001518; GB:BA000004; NID:gi0175792; PIDN:BA8069

A:Experimental source: strain C-125

C:Genetics:

A:Gene: murC

C:Superfamily: UDP-N-acetylmuramate-alanine ligase

Query Match

45.9%; Score 45; DB 2; Length 433;

Best Local Similarity 46.7%; Pred. No. 13; Mismatches 6; Indels 0; Gaps 0;

Matches 7; Conservative 2;

1 CFFRGGGFNPRYC 15

152 CEYRRHFLNRPDYC 166

RESULT 6

AD0809

xanthosine phosphorylase (EC 2.4.2.-) [imported] - Salmonella enterica subsp. enterica sero

C:Species: Salmonella enterica subsp. enterica serovar Typhi

A>Note: this species has also been called Salmonella typhi

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AD0809

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AD0809

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-277 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD07654.1; PID:gl6503641; GSPDB:GN00176

C:Genetics:

A:Gene: STY2658

C:Superfamily: purine-nucleoside phosphorylase

C:Keywords: glycosyltransferase; pentosyltransferase

Query Match

44.9%; Score 44; DB 2; Length 277;

Best Local Similarity 75.0%; Pred. No. 12; Mismatches 1; Indels 0; Gaps 0;

Matches 6; Conservative 1;

8 FNNHPRYC 15

6 FSHNPEYC 13

RESULT 7

JC7527

nuclear retroviral polymerase-like protein - rat

C;Species: *Rattus norvegicus* (Norway rat)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C;Accession: JC7527
 R;Graham, K.M.; Ko, C.; Park, K.; Sarge, K.; Park-Sarge, O.K.
 Biochem. Biophys. Res. Commun. 278, 48-57, 2000
 A;Title: Expression of an intracisternal A-particle-like element in rat ovary.
 A;Reference number: JC7527; MUID:20525406; PMID:11071854
 A;Contents: Ovary, granulosa cells
 A;Accession: JC7527
 A;Molecule type: mRNA
 A;Residues: 1-365 <GRA>
 A;Cross-references: GB:AA964260
 C;Comment: This protein is involved in cell proliferation, differentiation, folliculogenesis
 C;Genetics:
 A;Gene: iap-le
 C;Keywords: ovary; transformation

Query Match 44.9%; Score 44; DB 2; Length 365;
 Best Local Similarity 53.8%; Pred. No. 16;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 13
 | : | | | | | : |
 Db 276 CYDGEFDNHYPK 288

RESULT 8
 A43717
 alpha-galactosidase (EC 3.2.1.22), raffinose-specific - *Escherichia coli* plasmid D1021
 C;Species: *Escherichia coli*
 C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
 C;Accession: A43717; B35160
 R;Alanidis, C.; Schmid, K.; Schmitt, R.
 J. Bacteriol. 171, 6753-6763, 1989
 A;Title: Nucleotide sequences and operon structure of plasmid-borne genes mediating uptake
 A;Reference number: A43717; MUID:90078124; PMID:2556373
 A;Accession: A43717
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-708 <ASI>
 A;Cross-references: UNIPROT:P16551; GB:M27273; NID:gl47504; PIDN:AAA24497.1; PID:gl47505
 R;Alanidis, C.; Schmitt, R.
 J. Bacteriol. 172, 2178-2180, 1990
 A;Title: Regulatory elements of the raffinose operon: nucleotide sequences of operator and
 A;Reference number: A35160; MUID:90202743; PMID:2180920
 A;Accession: B35160
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-6 <AS2>
 A;Cross-references: GB:M29849; NID:gl47508; PIDN:AAA24501.1; PID:g551828
 C;Genetics:
 A;Gene: rafa
 A;Genome: plasmid
 C;Complex: homotetramer
 C;Function:
 A;Description: catalyzes the hydrolysis of raffinose to galactose and sucrose
 C;Superfamily: raffinose-specific alpha-galactosidase
 C;Keywords: glycoprotein; glycosidase; hydrolase

Query Match 44.9%; Score 44; DB 1; Length 708;
 Best Local Similarity 58.3%; Pred. No. 31;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 FRGGFFNNHNP 14
 : | : | | | | |
 Db 299 WEGIFYNNHNP 310

RESULT 9
 T13389
 hypothetical protein 115C2.10 - fruit fly (*Drosophila melanogaster*)
 C;Species: *Drosophila melanogaster*
 C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004

C;Accession: T13389
 R;Salles, C.; Valenti, P.; Darlamitsou, A.; Henderson, N.; Campbell, L.; Glover, D.
 submitted to the EMBL Data Library, May 1999
 A;Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.
 A;Reference number: Z17665
 A;Accession: T13389
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-1291 <CAT>
 A;Cross-references: UNIPROT:O77261; EMBL:AL031581; NID:el320978; PID:el426292; PIDN:CAA2
 C;Genetics:
 A;Cross-references: FlyBase:FBgn0020381
 A;Map position: X
 A;Introns: 238/3; 1225/1
 A;Note: EG:115C2.10

Query Match 44.9%; Score 44; DB 2; Length 1291;
 Best Local Similarity 46.7%; Pred. No. 57;
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 15
 | : | | | | | : |
 Db 363 CFYGEDFFGDSNRYC 377

RESULT 10
 S46106
 Probable membrane protein YBR230c - yeast (*Saccharomyces cerevisiae*)
 A;Alternate names: hypothetical protein YBR1527
 C;Species: *Saccharomyces cerevisiae*
 C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
 C;Accession: S46106
 R;Dubois, E.; El Bakkoury, M.; Glansdorff, N.; Messenguy, F.; Pierard, A.; Scherens, B.;
 submitted to the Protein Sequence Database, August 1994
 A;Reference number: S45782
 A;Accession: S46106
 A;Molecule type: DNA
 A;Residues: 1-134 <DUB>
 A;Cross-references: UNIPROT:P38325; EMBL:Z36099; NID:g536627; PID:g536628; GSPDB:GN00002
 A;Experimental source: strain S288C
 C;Genetics:
 A;Gene: MIPS:YBR230C
 A;Cross-references: SGD:S0000434
 A;Map position: 2R
 A;Introns: 4/2
 C;Superfamily: *Saccharomyces* probable membrane protein YBR230c
 C;Keywords: transmembrane protein
 F;70-88/Domain: transmembrane #status predicted <TM1>
 F;107-123/Domain: transmembrane #status predicted <TM2>

Query Match 43.9%; Score 43; DB 2; Length 134;
 Best Local Similarity 42.9%; Pred. No. 8.5;
 Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 14
 | : | : | | | : |
 Db 84 CYLCNGYANNHNP 97

RESULT 11
 F86382
 hypothetical protein F4P7.14 - *Arabidopsis thaliana*
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C;Accession: F86382
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

```

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F86382
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-373 <STO>
A;Cross-references: GB:AE005172; NID:gl1067280; PIDN:AAG28808.1; GSPDB:GN00141
C;Genetics:
A;Map position: 1

Query Match 43.9%; Score 43; DB 2; Length 373;
Best Local Similarity 61.5%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 13
||| ||| |||
Db 121 CFFNGGESRLNPR 133

RESULT 12
AH2348
hypothetical protein all4343 [imported] - Nostoc sp. (strain PCC 7120)
A;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
J;Accession: AH2348
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A>Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2348
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-660 <KUR>
A;Cross-references: UNIPROT:Q8YB57; GB:BA000019; PIDN:BA076042.1; PID:gl17133479; GSPDB:G
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all4343

Query Match 43.9%; Score 43; DB 2; Length 660;
Best Local Similarity 70.0%; Pred. No. 43;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 GGFNNHNP 14
||| ||| |||
Db 387 GGFNNRNRY 396

RESULT 13
C96559
PSF19.3 [imported] - Arabidopsis thaliana
A;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
J;Accession: C96559
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: C96559
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-124 <STO>
A;Cross-references: UNIPROT:Q9ZU26; GB:AE005173; NID:g4220444; PIDN:AAD12671.1; GSPDB:GN
C;Genetics:

```

```

A;Gene: F5F19.3
A;Map position: 1

Query Match 42.9%; Score 42; DB 2; Length 124;
Best Local Similarity 66.7%; Pred. No. 11;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CFFRGGFNN 9
||| ||| |||
Db 86 CMFRGGWYN 94

RESULT 14
H82875
oligopeptide transport system permease protein UU559 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: H82875
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to Genbank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: H82875
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-439 <GLA>
A;Cross-references: GB:AE002154; GB:AF222894; NID:g6899557; PIDN:AAF30972.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: oppF; UU559
A;Genetic code: SGC3

Query Match 42.9%; Score 42; DB 2; Length 439;
Best Local Similarity 66.7%; Pred. No. 41;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 FFNNHPRYC 15
||| ||| |||
Db 10 FFNNHKKRFC 18

RESULT 15
A56175
adhesive plaque protein MgpF2 precursor - Mediterranean mussel
C;Species: Mytilus galloprovincialis (Mediterranean mussel)
C;Date: 27-Apr-1995 #sequence_revision 03-Oct-1995 #text_change 09-Jul-2004
C;Accession: A56175
R;Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.
J. Biol. Chem. 270, 6698-6701, 1995
A>Title: Mussel adhesive plaque protein gene is a novel member of epidermal growth facto
A;Reference number: A56175; MUID:95204464; PMID:7896812
A;Accession: A56175
A;Molecule type: mRNA
A;Residues: 1-473 <INO>
A;Cross-references: UNIPROT:Q25464; GB:D43794; NID:g602767; PIDN:BA007852.1; PID:dl00843
C;Keywords: duplication
F;1-17/Domain: signal sequence #status predicted <Sig>
F;387-419/Domain: EGF homology <EGF1>
F;429-460/Domain: EGF homology <EGF>
F;23,36,43,56,75,382,424,455,468,473/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #

Query Match 42.9%; Score 42; DB 2; Length 473;
Best Local Similarity 30.4%; Pred. No. 44;
Matches 7; Conservative 4; Mismatches 4; Indels 8; Gaps 1;

Qy 1 CFFRGGFNNH- - - - -PRYC 15
||| :||| :|||
Db 69 CFCGGYGYNCNLKNACKPNQC 91

Search completed: October 4, 2005, 22:17:47
Job time : 41 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 21:42:35 ; Search time 170 Seconds
(without alignments)
45.183 Million cell updates/sec

Title: US-10-633-423-2

Perfect score: 98

Sequence: 1 CFFRGGFNNPRYC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	53.5	54.6	453	2 P92013	P92013 caenorhabdi
2	50	51.0	272	2 O01981	O01981 caenorhabdi
3	50	51.0	407	2 Q95Y59	Q95Y59 caenorhabdi
4	50	51.0	447	2 Q17816	Q17816 caenorhabdi
5	50	51.0	484	2 Q65280	Q65280 caenorhabdi
6	50	51.0	511	2 O16512	O16512 caenorhabdi
7	47	48.0	338	2 Q9XZ29	Q9XZ29 megaselia a
8	47	48.0	441	2 Q962Y5	Q962Y5 ascaris suu
9	47	48.0	1362	2 Q9PV24	Q9PV24 xenopus lae
10	46	46.9	318	2 Q91X81	Q91X81 mus musculu
11	46	46.9	334	2 Q8BU67	Q8BU67 mus musculu
12	46	46.9	334	2 Q8AAE1	Q8AAE1 bacteroides
13	46	46.9	354	2 Q73LX3	Q73LX3 treponema d
14	46	46.9	384	2 Q9Y393	Q9Y393 homo sapien
15	46	46.9	393	2 Q9BUL0	Q9BUL0 homo sapien
16	46	46.9	394	2 Q8BUT7	Q8BUT7 mus musculu
17	46	46.9	765	1 EAR1_MOUSE	O70445 mus musculu
18	46	46.9	817	2 Q6DI74	Q6DI74 mus musculu
19	46	46.9	855	2 Q6GP17	Q6GP17 xenopus lae
20	46	46.9	860	2 Q8BUN0	Q8BUN0 mus musculu
21	46	46.9	874	2 Q6Q784	Q6Q784 mus musculu
22	46	46.9	1330	2 Q9SXX5	Q9SXX5 heterosigma
23	45	45.9	433	1 MURC_BACHD	Q9K7W1 bacillus ha
24	45	45.9	525	2 Q7PYN5	Q7PYN5 anopheles g
25	45	45.9	712	2 Q8T2K1	Q8T2K1 dictyosteli
26	45	45.9	740	2 Q88T11	Q88T11 lactobacill
27	44.5	45.4	368	2 Q6PBT4	Q6PBT4 brachydanio
28	44.5	45.4	3303	2 Q8I339	Q8I339 plasmodium
29	44	44.9	203	2 Q6LP28	Q6LP28 photobacter
30	44	44.9	277	2 Q8Z4W8	Q8Z4W8 salmonella
31	44	44.9	277	2 Q8ZN94	Q8ZN94 salmonella

32 44 44.9 320 2 Q75AJ0 Q75AJ0 ashbya goss
33 44 44.9 440 2 Q8MYV0 Q8MYV0 drosophila
34 44 44.9 440 2 Q9VSG4 Q9VSG4 drosophila
35 44 44.9 708 1 RAFA_ECOLI P16551 escherichia
36 44 44.9 782 2 Q8T029 Q8T029 drosophila
37 44 44.9 1111 2 Q7SAB0 Q7SAB0 neurospora
38 44 44.9 1291 2 Q77261 Q77261 drosophila
39 44 44.9 1300 2 Q9WSE0 Q9WSE0 drosophila
40 43 43.9 61 2 Q8FDI1 Q8FDI1 escherichia
41 43 43.9 134 1 YB80_YEAST P38325 saccharomyc
42 43 43.9 211 1 SSRP_TROW8 Q83nf6 tropheryma
43 43 43.9 211 1 SSRP_TROW8 Q83nf3 tropheryma
44 43 43.9 232 2 Q8NAJ2 Q8NAJ2 homo sapien
45 43 43.9 284 2 Q8KCK9 Q8KCK9 chlorobium

ALIGNMENTS

RESULT 1
P92013 PRELIMINARY; PRT; 453 AA.
AC P92013
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein R10D12.15 (Hypothetical protein T26F2.1).
GN ORFNames=R10D12.15, T26F2.1;
OS Caenorhabditis elegans.
OC Rukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX PERCY C.M.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX PERCY C.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z81109; CAB04842.1; JOINED.
DR EMBL; Z82054; CAB04842.1; -.
DR EMBL; Z82054; CAB04840.1; -.
DR EMBL; Z82054; CAB03255.1; JOINED.
DR EMBL; Z81109; CAB03255.1; -.
DR PIR; T24127; T24127.
DR HSSP; P23951; ICNS.
DR WormBase; WSGene00011196; R10D12.15.
DR WormBase; WSGene00012057; T26F2.1.
DR WormPep; R10D12.15; CE12700.
DR WormPep; T26F2.1; CE12700.
DR GO; GO:0004568; P:chitinase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00182; Glyco_hydro_19; 1.
KW Hypothetical protein.
SQ SEQUENCE 453 AA; 50165 MW; 6711081547BE6D54 CRC64;

Query Match 54.6%; Score 53.5; DB 2; Length 453;
Best Local Similarity 28.6%; Pred.No.3.4;
Matches 12; Conservative 3; Mismatches 0; Indels 27; Gaps 2;

Qy 1 CFFRGGFN-----HNP-----RYC 15
 ||:||||:
 Db 160 CFYRGGFNPEGGSSNFLNPFPGHSPDTGNSCTGAGRYC 201

RESULT 2

001981 PRELIMINARY; PRT; 272 AA.
 ID 001981
 AC 001981;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein CS1E3.8.
 GN ORFNames=CS1E3.8;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode C.elegans: A platform for
 investigating biology."; Science 282:2012-2018(1998).
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilkinson J.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z78410; CAB01643.1; -;
 DR PIR; T20139; T20139.
 DR WormBase; WBGene00008253; CS1E3.8.
 DR WormPep; CS1E3.8; CE08942.
 KW Hypothetical protein.
 SQ SEQUENCE 272 AA; 30910 MW; AE945211D9935DS7 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 272;
 Best Local Similarity 77.8%; Pred. No. 7.6;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
 ||:||||:
 Db 157 CFYRGGFN 165

RESULT 3

Q95Y59 PRELIMINARY; PRT; 407 AA.
 ID Q95Y59
 AC Q95Y59;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein Y50D4A.3.
 GN Name=Y50D4A.3; ORFNames=Y50D4A.3;
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RG WormBase Consortium;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Lamar B., Le T., Elliott G.;

RT "The sequence of C. elegans cosmid Y50D4A."; Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 RN [12]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RG WormBase Consortium;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC006794; AAK68504.1; -;
 DR HSSP; P23951; 2BAA.
 DR WormBase; WBGene00021737; Y50D4A.3.
 DR WormPep; Y50D4A.3; CE27638.
 DR GO; GO:0004568; F:chitinase activity; IEA.
 DR GO; GO:0016998; P:cell wall catabolism; IEA.
 DR GO; GO:0006032; P:chitin catabolism; IEA.
 DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
 DR InterPro; IPR000726; Glyco_hydro_19.
 DR Pfam; PF00182; Glyco_hydro_19; 1.
 DR ProDom; PD35490; Glyco_hydro_19; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 407 AA; 45323 MW; 3B6640DC4918A352 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 407;
 Best Local Similarity 77.8%; Pred. No. 11;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
 ||:||||:
 Db 117 CFYRGGFN 125

```
RESULT 4
Q17816 PRELIMINARY; PRT; 447 AA.
AC Q17816
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein C08B6.4a.
GN Name=C08B6.4a; ORFNames=C08B6.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilkinson J.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z72502; CA95587.1; -.
DR PIR; T19078; T19078.
DR HSP; P23951; 2BAA.
DR WormBase; WBGene00007425; C08B6.4.
DR WormPep; C08B6.4; C505241.
DR GO; GO:0004568; F:chitinase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR Hypothetical protein.
KW SEQUENCE 447 AA; 49583 MW; 684EC67A9B13D17 CRC64;
SQ SEQUENCE 447 AA; 49583 MW; 684EC67A9B13D17 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 447;
Best Local Similarity 77.8%; Pred. No. 12;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
Db 154 CFYRGGFN 162

RESULT 5
Q652B0 PRELIMINARY; PRT; 484 AA.
AC Q652B0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein C08B6.4b.
GN Name=C08B6.4b;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RL Science 282:2012-2018(1998).
RN [2]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilkinson J.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z72502; CAH19081.1; -.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD354900; Glyco_hydro_19; 1.
KW Hypothetical protein.
SQ SEQUENCE 484 AA; 53734 MW; E1EDF5E8CFB84A46 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 484;
Best Local Similarity 77.8%; Pred. No. 13;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
Db 191 CFYRGGFN 199

RESULT 6
Q16512 PRELIMINARY; PRT; 511 AA.
AC Q16512;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein T05H4.7.
GN Name=T05H4.7; ORFNames=T05H4.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Blanchard M.;
RT "The sequence of C. elegans cosmid T05H4.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF016452; AAB66013.1; -.
DR PIR; T31899; T31899.
DR HSP; P23951; 2BAA.
DR WormBase; WBGene00020270; T05H4.7.
DR WormPep; T05H4.7; CE13281.
DR GO; GO:0004568; F:chitinase activity; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR ProDom; PD354900; Glyco_hydro_19; 1.
DR Hypothetical protein.
KW SEQUENCE 511 AA; 57956 MW; 97CSBF697F3BC7A0 CRC64;
SQ SEQUENCE 511 AA; 57956 MW; 97CSBF697F3BC7A0 CRC64;

Query Match 51.0%; Score 50; DB 2; Length 511;
```

Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
Db 204 CFYRGGFN 212

RESULT 7

ID Q9XZ29 PRELIMINARY; PRT; 338 AA.
AC Q9XZ29;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bicoïd protein (Fragment).
GN Name:bcd;
OS Megaselia abdita.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Platypezoidae; Phoridae; Megaseliini; Megaselia.
OX NCBI_TaxID=88686;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9919262; PubMed=10097115; DOI=10.1073/pnas.96.7.3786;
RA Stauber M., Jaekle H., Schmidt-Ott U.;
RT "The anterior determinant bicoïd of Drosophila is a derived Hox class 3 gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:3786-3789(1999).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
DR EMBL; AJ133024; CAB40892.1; -.
DR HSP; P13297; IIG7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeobox_like.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00771; HOMEBOX 2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
FT NON_TER 338 338
SQ SEQUENCE 338 AA; 37818 MW; D50D16E7096E7DE1 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 338;
Best Local Similarity 72.7%; Pred. No. 28;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 11
Db 205 CFASGGFFN 215

RESULT 8

ID Q962Y5 PRELIMINARY; PRT; 441 AA.
AC Q962Y5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative chitinase (EC 3.2.1.14) (Fragment).
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RA Geng J., Komuniecki P., Komuniecki R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF395836; AAK93964.1; -.
DR HSP; P23951; 1CNS.
DR GO; GO:0004568; F:chitinase activity; IEA.

DR GO; GO:0016798; F:hydrolase activity, acting on glycosyl bonds; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR GO; GO:0006032; P:chitin catabolism; IEA.
DR GO; GO:0009613; P:response to pest, pathogen or parasite; IEA.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00182; Glyco_hydro_19; 1.
KW Glycosidase; Hydrolase.
FT NON_TER 1 1
SQ SEQUENCE 441 AA; 49416 MW; C1EF9BE91D2B65B1 CRC64;

Query Match 48.0%; Score 47; DB 2; Length 441;
Best Local Similarity 66.7%; Pred. No. 37;

Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CFFRGGFN 9
Db 140 CYRGGGFN 148

RESULT 9

ID Q9PVZ4 PRELIMINARY; PRT; 1362 AA.
AC Q9PVZ4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Insulin receptor precursor.
GN Name:insulin receptor;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Groigno L., Richard-Parpailion L., Boujard D.;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin receptor subfamily.
DR EMBL; AJ132556; CAB46565.1; -.
DR PIR; B41122; B41122.
DR HSP; P06213; I144.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0005006; F:epidermal growth factor receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000494; EGFR_L.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR006211; Furin-like.
DR InterPro; IPR006212; Furin repeat.
DR InterPro; IPR009030; Growth factor receptor.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Protein kinase.
DR InterPro; IPR002011; RecepttyrkinasII.
DR InterPro; IPR001245; Tyr_kinase.
DR InterPro; IPR008266; Tyr_kinase_AS.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00757; Furin-like; 1.
DR Pfam; PF01030; Recept_L domain; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 3.
DR SMART; SM00261; FU; 2.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00853; FN3; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

DR PROSITE, PS50011; PROTEIN KINASE DOM; 1.
DR PROSITE, PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE, PS00239; RECEPTOR TYR KIN II; 1.
KW ATP-binding; Kinase; Phosphorylation; Receptor; Signal; Transferase;
FT Transmembrane; Tyrosine-protein kinase.
FT SIGNAL 1 37 Potential.
FT CHAIN 759 1362 insulin receptor, beta-subunit.
FT CHAIN 38 754 insulin receptor, alpha-subunit.
SQ SEQUENCE 1362 AA; 153755 MW; 7B8BF2FB7EFDA01B CRC64;

Query Match 48.0%; Score 47; DB 2; Length 1362;

Best Local Similarity 77.8%; Pred. NO. 1.1e+02; Mismatches 1; Indels 0; Gaps 0;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 7 FFHNPVRYC 15

Db 464 FFHNPRLC 472

RESULT 10

Q91X81 PRELIMINARY; PRT; 318 AA.
AC Q91X81
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Suv420h1 protein.
CN Name=Suv420h1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Stapchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Stappert M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Rodriguez A.C., Touchman J.W., Green E.D., Dickson M.C.,
RA Krzywinski M.I., Skalska U., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

SEQUENCE FROM N.A.

STRAIN=FVB/N; TISSUE=Salivary gland;

Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

-1- SIMILARITY: Contains 1 SET domain.

EMBL: BC011214; AAH11214.1; --

MGB: MGI:2444557; Suv420h1.

GO: GO:0000780; C:condensed nuclear chromosome, pericentric r. .; IDA.

GO: GO:0042799; P:histone lysine N-methyltransferase activity. .; IDA.

GO: GO:0016571; P:histone methylation; IDA.

Pfam: PF00856; SET; 1.

SMART: SM00317; SET; 1.

PROSITE: PS50280; SET; 1.

SEQUENCE 318 AA; 36487 MW; 7B69279AD9E3AB61 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 318;

Best Local Similarity 40.0%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNPVC 15

Db 297 CYGDGFGFNNPVC 311

RESULT 11

Q8BU67 PRELIMINARY; PRT; 318 AA.
AC Q8BU67
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 0 day neonate lung cDNA. RIKEN full-length enriched
DE library, clone: E030040N07 product: SIMILAR TO CGI-85 PROTEIN
DE homolog.
DE Name=Suv420h1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA the FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20550913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsu T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

```

RA Hayaashida K., Hayateu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ichii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyma T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Saeki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tegami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tonari A., Toya T., Yasunishi A., Muramatsu M., Hayaashizaki Y.,
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 SET domain.
DR EMBL: AK087267; BAC39834.1; -.
DR MGD: MGI:2444557; Suv420hl.
DR GO: GO:0000780; C:condensed nuclear chromosome, pericentric r. . ; IDA.
DR GO: GO:0042799; P:histone lysine N-methyltransferase activity. . ; IDA.
DR GO: GO:0016571; P:histone methylation; IDA.
DR InterPro: IPR001214; SET.
DR Pfam: PF00856; SET; 1.
DR SMART: SM00317; SET; 1.
DR PROSITE: PS50280; SET; 1.
DR SEQUENCE 318 AA; 36500 MW; 69CD8984CC084CA6 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 318;
Best Local Similarity 40.0%; Pred. No. 38;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGFFNNHNPYC 15
|::|||::|
Db 297 CYGDFGFFGENNEFC 311

RESULT 12
Q8AAAF1 PRELIMINARY; PRT; 334 AA.
ID Q8AAAF1
AC Q8AAAF1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT0514;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL: AE016928; AA075621.1; -.
KW Complete proteome.
SQ SEQUENCE 334 AA; 39493 MW; B4B36FBFA9CE2F09 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 334;
Best Local Similarity 66.7%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 RGGFFNNHNPYC 15
|::|||::|
Db 71 RGSFFNGNPMTC 82

RESULT 13
Q73LX3 PRELIMINARY; PRT; 354 AA.
ID Q73LX3
AC Q73LX3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE FemAB family protein.

```

```

GN OrderedLocusNames=TDE1739;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidsen T.M., DeBoy R.T., Fouts D.E., Haft D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Gbregregoris E., Geer K., Tsagias G., Malek J.A., Ayodeji B.,
RA Shatman S., McLeod M.P., Smajs D., Howell J.K., Pal S., Amin A.,
RA Washioth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Wainstock G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL: AE017251; AAS12254.1; -.
DR TIGR: TDE1739; -.
DR GO: GO:0009252; P:peptidoglycan biosynthesis; IEA.
DR InterPro: IPR003447; Meth_resist.
DR Pfam: PF02388; FemAB; 1.
KW Complete proteome.
SQ SEQUENCE 354 AA; 41498 MW; 271DBBB50DB47095 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 354;
Best Local Similarity 72.7%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RGGFFNNHNPYC 14
|::|||::|
Db 206 RGGFFNNHNGEY 216

RESULT 14
Q9Y393 PRELIMINARY; PRT; 384 AA.
ID Q9Y393
AC Q9Y393;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CGI-85 protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20272150; PubMed=10810093; DOI=10.1101/gr.10.5.703;
RA Lai C.-H., Chou C.-Y., Ch'ang L.-Y., Liu C.-S., Lin W.-C.;
RT "Identification of novel human genes evolutionarily conserved in
Caenorhabditis elegans by comparative proteomics.";
RL Genome Res. 10:703-713(2000).
CC -1- SIMILARITY: Contains 1 SET domain.
DR EMBL: AF151843; AAD34080.1; -.
DR InterPro: IPR001214; SET.
DR Pfam: PF00856; SET; 1.
DR SMART: SM00317; SET; 1.
DR PROSITE: PS50280; SET; 1.
DR SEQUENCE 384 AA; 43439 MW; 88B3BC0C5D25DB67 CRC64;

Query Match 46.9%; Score 46; DB 2; Length 384;
Best Local Similarity 40.0%; Pred. No. 46;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGFFNNHNPYC 15
|::|||::|
Db 296 CYGDFGFFGENNEFC 310

RESULT 15

```

Q9BULO
ID Q9BULO PRELIMINARY; PRT; 393 AA.
AC Q9BULO; Q6P150;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Suppressor of variegation 4-20 homolog 1, isoform 2.
GN Name-SUV420H1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX Director MGC Project;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX Director MGC Project;
RA Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 SET domain.
DR EMBL; BC002522; AAH02522.2; -.
DR EMBL; BC065287; AAH65287.1; -.
DR InterPro; IPR001214; SET.
DR Pfam; PF00856; SET; 1.
DR SMART; SM00317; SET; 1.
DR PROSITE; PS00280; SET; 1.
SQ SEQUENCE 393 AA; 44618 MW; A9690C2164F3384E CRC64;

Query Match 46.9%; Score 46; DB 2; Length 393;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGFFFNPNRYC 15
|::||| |:
Db 305 CYVGDFGFGENEFC 319

Search completed: October 4, 2005, 22:15:08
Job time : 173 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2005, 22:03:45 ; Search time 42 Seconds
(without alignments)
26.660 Million cell updates/sec

Title: US-10-633-423-2

Perfect score: 98

Sequence: 1 CFFRGGFFNNPRYC 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*

5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pcp.*

6: /cgn2_6/ptodata/1/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	46.9	876	4	US-09-764-176-10
2	42	42.9	315	4	US-09-187-906-19
3	42	42.9	400	3	US-09-220-528-63
4	42	42.9	400	4	US-09-187-906-21
5	42	42.9	400	4	US-09-949-016-9079
6	41	41.8	60	4	US-09-248-796A-27332
7	41	41.8	350	4	US-09-270-767-43557
8	41	41.8	521	2	US-08-504-048-9
9	41	41.8	653	4	US-09-171-937C-27
10	41	41.8	779	4	US-09-171-937C-25
11	40	40.8	495	4	US-09-489-039A-10528
12	40	40.8	588	4	US-09-252-991A-18578
13	40	40.8	713	2	US-08-849-212-4
14	40	40.8	832	4	US-09-489-039A-12438
15	39.5	40.3	313	2	US-08-990-379-7
16	39.5	40.3	314	4	US-09-164-193-22
17	39.5	40.3	314	4	US-09-221-448A-22
18	39.5	40.3	458	4	US-09-543-681A-6324
19	39	39.8	15	6	5227466-7
20	39	39.8	15	6	5227466-7
21	39	39.8	267	4	US-09-634-238-288
22	39	39.8	379	4	US-09-543-681A-7297
23	39	39.8	383	3	US-08-857-076-105
24	38.5	39.3	368	4	US-09-248-796A-16800
25	38.5	39.3	527	3	US-08-369-822C-25
26	38.5	39.3	527	3	US-08-582-776C-40
27	38.5	39.3	527	3	US-08-434-831B-37

28	38	38.8	40	1	US-08-144-121-11	Sequence 11, Appl
29	38	38.8	40	2	US-08-735-893-11	Sequence 11, Appl
30	38	38.8	71	4	US-09-248-796A-26506	Sequence 26506, A
31	38	38.8	78	4	US-09-248-796A-24372	Sequence 24372, A
32	38	38.8	81	4	US-09-543-681A-8169	Sequence 8169, Ap
33	38	38.8	109	4	US-09-583-110-4499	Sequence 4499, Ap
34	38	38.8	160	4	US-09-489-039A-8564	Sequence 8564, Ap
35	38	38.8	169	4	US-09-107-433-2681	Sequence 2681, Ap
36	38	38.8	196	4	US-09-270-767-34969	Sequence 34969, A
37	38	38.8	196	4	US-09-270-767-50186	Sequence 50186, A
38	38	38.8	283	4	US-09-248-796A-16847	Sequence 16847, A
39	38	38.8	316	4	US-09-489-039A-12256	Sequence 12256, A
40	38	38.8	335	4	US-09-583-110-5099	Sequence 5099, Ap
41	38	38.8	370	3	US-08-857-076-104	Sequence 104, App
42	38	38.8	377	4	US-09-902-540-13663	Sequence 13663, A
43	38	38.8	447	4	US-09-328-352-7917	Sequence 7917, Ap
44	38	38.8	456	4	US-09-489-039A-8419	Sequence 8419, Ap
45	38	38.8	462	2	US-08-477-451-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-09-764-176-10

; Sequence 10, Application US/09764176

; Patent No. 6809189

; GENERAL INFORMATION:

; APPLICANT: NOTEBORN, Mathieu Hubertus Maria

; APPLICANT: DANEN-VAN OORSCHOT, Astrid Adriana Anna Maria

; APPLICANT: ROHN, Jennifer Leigh

; APPLICANT: WEISS, Bertram

; TITLE OF INVENTION: APOPTIN-ASSOCIATING PROTEIN

; FILE REFERENCE: 4725US

; CURRENT APPLICATION NUMBER: US/09/764,176

; CURRENT FILING DATE: 2001-01-17

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 10

; LENGTH: 876

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Description of Sequence: Amino acid sequence deduced from the nuc

; OTHER INFORMATION: leic acid sequence of AAP-

US-09-764-176-10

Query Match 46.9%; Score 46; DB 4; Length 876;

Best Local Similarity 40.0%; Pred. No. 45;

Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CFFRGGFFNNPRYC 15

::: ||| |::

Db 296 CYYGDDFFGNNRFC 310

RESULT 2

US-09-187-906-19

; Sequence 19, Application US/09187906

; Patent No. 6677135

; GENERAL INFORMATION:

; APPLICANT: BIOGEN, INC.

; TITLE OF INVENTION: Ret Ligand (RetL) for Stimulating Neural

; TITLE OF INVENTION: and Renal Growth

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Biogen, Inc.

; STREET: 14 Cambridge Center

; CITY: Cambridge

; STATE: MA

; COUNTRY: USA

; ZIP: 02142

Query Match	Best Local Similarity	Score	DB	Length	400;
Query Match	42.9%;	Score 42;	DB 4;	Length 400;	
Best Local Similarity	58.3%;	Pred. No. 84;			
Matches	7; Conservative	1; Mismatches	4; Indels	0; Gaps	0;
QY	1	CFFRGGFFHNHP 12			
DB	328	CEMLEGFFSHNP 339			
RESULT 3					
US-09-220-528-63					
Sequence 63, Application US/09220528A					
Patent No. 6284540					
GENERAL INFORMATION:					
APPLICANT: Milbrandt, Jeffrey D.					
TITLE OF INVENTION: Artemin, A No. 6284540el Neurotrophic Factor					
FILE REFERENCE: 6029-7998					
CURRENT APPLICATION NUMBER: US/09/220,528A					
CURRENT FILING DATE: 1998-12-24					
EARLIER APPLICATION NUMBER: 09/218,698					
EARLIER FILING DATE: 1998-12-22					
EARLIER APPLICATION NUMBER: 60/108,148					
EARLIER FILING DATE: 1998-11-12					
EARLIER APPLICATION NUMBER: 09/163,283					
EARLIER FILING DATE: 1998-09-29					
NUMBER OF SEQ ID NOS: 120					
SOFTWARE: patentin Ver. 2.0					
SEQ ID NO 63					
LENGTH: 400					
TYPE: PRT					
ORGANISM: Homo sapiens					
US-09-220-528-63					
Query Match	42.9%;	Score 42;	DB 3;	Length 400;	
Best Local Similarity	58.3%;	Pred. No. 66;			
Matches	7; Conservative	1; Mismatches	4; Indels	0; Gaps	0;
QY	1	CFFRGGFFHNHP 12			
DB	243	CEMLEGFFSHNP 254			
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: PatentIn Release #1.0, Version #1.30					
CURRENT APPLICATION DATA:					
APPLICATION NUMBER: US/09/187,906					
FILING DATE:					
CLASSIFICATION:					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: PCT/US97/07726					
FILING DATE: 07-MAY-97					
APPLICATION NUMBER: US 60/017,427					
FILING DATE: 08-MAY-96					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: US 60/019,300					
FILING DATE: 07-JUN-96					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: US 60/021,859					
FILING DATE: 10-APR-97					
PRIOR APPLICATION DATA:					
APPLICATION NUMBER: US 60/043,533					
FILING DATE: 10-APR-97					
ATTORNEY/AGENT INFORMATION:					
NAME: Kaplan, Warren A.					
REGISTRATION NUMBER: 34,199					
REFERENCE/DOCKET NUMBER: A008 PCT CIP					
TELECOMMUNICATION INFORMATION:					
TELEPHONE: 617-679-2400					
TELEFAX: 617-679-2838					
INFORMATION FOR SEQ ID NO: 19:					
SEQUENCE CHARACTERISTICS:					
LENGTH: 315 amino acids					
TYPE: amino acid					
TOPOLOGY: linear					
MOLECULE TYPE: protein					
US-09-187-906-19					
Query Match	42.9%;	Score 42;	DB 4;	Length 315;	
Best Local Similarity	58.3%;	Pred. No. 66;			
Matches	7; Conservative	1; Mismatches	4; Indels	0; Gaps	0;
QY	1	CFFRGGFFHNHP 12			
DB	243	CEMLEGFFSHNP 254			
COMPUTER READABLE FORM:					
MEDIUM TYPE: Floppy disk					
COMPUTER: IBM PC compatible					
OPERATING SYSTEM: PC-DOS/MS-DOS					
SOFTWARE: PatentIn Release #1.0, Version #1.30					
CURRENT APPLICATION DATA:					

```
; Sequence 9079, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9079
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9079

Query Match          42.9%; Score 42; DB 4; Length 400;
Best Local Similarity 58.3%; Pred. No. 84;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHP 12
Db 328 CEMLEGFNSHP 339

RESULT 6
US-09-248-796A-27332
; Sequence 27332, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR FILING DATE: 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27332
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27332

Query Match          41.8%; Score 41; DB 4; Length 60;
Best Local Similarity 63.6%; Pred. No. 17;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 5 GGFNNHPRYC 15
Db 47 GGCFYVPRYC 57

RESULT 7
US-09-270-767-43557
; Sequence 43557, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
```

```
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43557
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-43557

Query Match          41.8%; Score 41; DB 4; Length 350;
Best Local Similarity 60.0%; Pred. No. 11e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 FFRGGFFNNH 11
Db 259 FYRGGYFREN 268

RESULT 8
US-08-504-048-9
; Sequence 9, Application US/08504048
; Patent No. 5843674
; GENERAL INFORMATION:
; APPLICANT: TAKIMOTO, Hiroyuki
; APPLICANT: SUZUKI, Satoshi
; APPLICANT: SHIBATA, Koushi
; APPLICANT: MASUI, Shigeki
; TITLE OF INVENTION: ANTI-HUMAN TYROSINASE
; TITLE OF INVENTION: MONOCLONAL ANTIBODY
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/504,048
; FILING DATE:
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: SATOT27.001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-504-048-9

Query Match          41.8%; Score 41; DB 2; Length 521;
Best Local Similarity 46.7%; Pred. No. 1.6e+02;
Matches 7; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

Qy 1 CFFRGGFNNHNP 15
Db 99 CHCNGFSGHNGCYC 113

RESULT 9

US-09-171-937C-27
; Sequence 27, Application US/09171937C
; Patent No. 6833490
; GENERAL INFORMATION:
; APPLICANT: GODDIJN, Oscar Johannes Maria
; PEN, Jan
; SNEEKENS, Josephus Christianus M.
; TITLE OF INVENTION: Regulating metabolism by modifying the
; level of trehalose-6-phosphate
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61 STREET
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-Dos/MS-Dos
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,937C
; FILING DATE: 28-Apr-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02497
; FILING DATE: 02-MAY-1997
; APPLICATION NUMBER: EP 96.201.225.8
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: EP 96.202.128.3
; FILING DATE: 26-JUL-1996
; APPLICATION NUMBER: EP 96.202.395.8
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MASS, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-011967-1
; TELEPHONE: (212) 708-1890
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 653 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:
US-09-171-937C-27

Query Match 41.8%; Score 41; DB 4; Length 653;
Best Local Similarity 58.3%; Pred. No. 2e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 12
Db 130 CRFKIGFHLSP 141

RESULT 10

US-09-171-937C-25
; Sequence 25, Application US/09171937C
; Patent No. 6833490
; GENERAL INFORMATION:
; APPLICANT: GODDIJN, Oscar Johannes Maria
; PEN, Jan

; SNEEKENS, Josephus Christianus M.
; TITLE OF INVENTION: Regulating metabolism by modifying the
; level of trehalose-6-phosphate
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LADAS & PARRY
; STREET: 26 WEST 61 STREET
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-Dos/MS-Dos
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,937C
; FILING DATE: 28-Apr-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP97/02497
; FILING DATE: 02-MAY-1997
; APPLICATION NUMBER: EP 96.201.225.8
; FILING DATE: 03-MAY-1996
; APPLICATION NUMBER: EP 96.202.128.3
; FILING DATE: 26-JUL-1996
; APPLICATION NUMBER: EP 96.202.395.8
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: MASS, Clifford J.
; REGISTRATION NUMBER: 30,086
; REFERENCE/DOCKET NUMBER: U-011967-1
; TELEPHONE: (212) 708-1890
; TELEFAX: (212) 246-8959
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 779 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-171-937C-25

Query Match 41.8%; Score 41; DB 4; Length 779;
Best Local Similarity 58.3%; Pred. No. 2.4e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 12
Db 130 CRFKIGFHLSP 141

RESULT 11

US-09-489-039A-10528
; Sequence 10528, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gaty Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10528
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10528

```
Query Match      40.8%; Score 40; DB 4; Length 495;
Best Local Similarity 53.8%; Pred. No. 2.2e+02;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CFFRGGFNNHNP 13
   |||||
Db 403 CFFTTGFEVFGPQ 415

RESULT 12
US-09-252-991A-18578
; Sequence 18578, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18578
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18578

Query Match      40.8%; Score 40; DB 4; Length 588;
Best Local Similarity 54.5%; Pred. No. 2.6e+02;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RGGFFNNHNP 14
   |||||
Db 474 RGGFAGHGPKH 484

RESULT 13
US-08-849-212-4
; Sequence 4, Application US/08849212
; Patent No. 5827698
; GENERAL INFORMATION:
; APPLICANT: KIKUCHI, YOSHIMI
; APPLICANT: SUZUKI, TOMOKO
; APPLICANT: KOJIMA, HIROYUKI
; TITLE OF INVENTION: NOVEL LYSINE DECARBOXYLASE GENE AND
; METHOD OF PRODUCING L-LYSINE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESS: P.C.
; STREET: 1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/849,212
; FILING DATE: 09-JUN-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6/306386
; FILING DATE: 09-DEC-1994
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-856-OPCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 713 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-849-212-4

Query Match      40.8%; Score 40; DB 2; Length 713;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 FFNNHPRYC 15
   | |||||
Db 45 FIEHNPRIC 53

RESULT 14
US-09-489-039A-12438
; Sequence 12438, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12438
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12438

Query Match      40.8%; Score 40; DB 4; Length 832;
Best Local Similarity 66.7%; Pred. No. 3.7e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 FFNNHPRYC 15
   | |||||
Db 160 FIEHNPRIC 168

RESULT 15
US-08-990-379-7
; Sequence 7, Application US/08990379
; Patent No. 5998188
; GENERAL INFORMATION:
; APPLICANT: Stork, Philip J
; APPLICANT: Misra-Presse, Anita
; TITLE OF INVENTION: Mitogen Activated Protein Kinase Phosphatase cDNAs and
; THEIR BIOLOGICALLY ACTIVE EXPRESSION PRODUCTS
; FILE REFERENCE: 4104-000322USA
; CURRENT APPLICATION NUMBER: US/08/990,379
; CURRENT FILING DATE: 1997-12-15
; EARLIER APPLICATION NUMBER: PCT/US96/10402
; EARLIER FILING DATE: 1996-06-14
; EARLIER APPLICATION NUMBER: 60/000,263
; EARLIER FILING DATE: 1995-06-16
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 313
; TYPE: PRT
```

i ORGANISM: Mus sp.
US-08-990-379-7

Query Match 40.3%; Score 39.5; DB 2; Length 313;
Best Local Similarity 53.3%; Pred. No. 1.6e+02;
Matches 8; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

Qy 1 CFFRGGFNNHNPYC 15
||| : ||
Db 125 CFLRGPKSQ-TYC 138

Search completed: October 4, 2005, 22:18:35
Job time : 43 secs


```

; SEQ ID NO 191962
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_106656C.1.pep
US-10-425-115-191962

Query Match      51.5%; Score 50.5; DB 16; Length 267;
Best Local Similarity 60.0%; Pred. No. 8.4;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CFFRGGFNNPRYC 15
Db 73 CFFK-GLFNCSPRYC 86

RESULT 7
US-10-425-114-36595
; Sequence 36595, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 36595
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3076-031-H5_FLI.pep
US-10-425-114-36595

Query Match      51.5%; Score 50.5; DB 15; Length 293;
Best Local Similarity 60.0%; Pred. No. 9.2;
Matches 9; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

QY 1 CFFRGGFNNPRYC 15
Db 99 CFFK-GLFNCSPRYC 112

RESULT 8
US-10-425-115-191953
; Sequence 191953, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 191953
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_106647C.1.pep
US-10-425-115-191953

```

Db 204 CFYRGGFYN 212

RESULT 11

US-10-425-115-339901
; Sequence 339901, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 339901
; LENGTH: 124

; TYPE: PRT
; ORGANISM: Zea mays

; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_73158C.1.pep

US-10-425-115-339901

Query Match 50.0%; Score 49; DB 16; Length 124;

Best Local Similarity 66.7%; Pred. No. 6.8;

Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RGGFFNNHNPYC 15

|||||:|

Db 86 RGGFFNMTPKRC 97

RESULT 12

US-10-425-115-344097

; Sequence 344097, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 344097
; LENGTH: 109

; TYPE: PRT
; ORGANISM: Zea mays

; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_76978C.1.pep

US-10-425-115-344097

Query Match 48.0%; Score 47; DB 16; Length 109;

Best Local Similarity 52.9%; Pred. No. 12;

Matches 9; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

Qy 1 CFFRG--GFFNNHNPYC 15

|||||:|

Db 40 CFFSGKRAFFLASPRFC 56

RESULT 13

US-10-425-115-290598

; Sequence 290598, Application US/10425115
; Publication No. US20040214272A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 290598
; LENGTH: 113
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(113)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_28114C.1.pep
US-10-425-115-290598

Query Match 48.0%; Score 47; DB 16; Length 113;

Best Local Similarity 58.3%; Pred. No. 13;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 4 RGGFFNNHNPYC 15

|||||:|

Db 7 KGGFFMHSPSVC 18

RESULT 14

US-10-424-599-251796

; Sequence 251796, Application US/10424599
; Publication No. US20040031072A1

; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J

; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 251796
; LENGTH: 114

; TYPE: PRT
; ORGANISM: Glycine max

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6939C.1.pep

US-10-424-599-251796

Query Match 48.0%; Score 47; DB 15; Length 114;

Best Local Similarity 75.0%; Pred. No. 13;

Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 FFRGGFFNNHNP 13

|||||

Db 67 FTRGGGNNHNP 78

RESULT 15

US-09-764-861-42

; Sequence 42, Application US/09764861
; Publication No. US20020086811A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT204

; CURRENT APPLICATION NUMBER: US/09/764,861
; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

```
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (157)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-861-42

Query Match      46.9%; Score 46; DB 9; Length 166;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY      1 CFFRGGFFNNPRYC 15
Db      114 CYGCGFFGENNEFC 128
      |::||| |
      |::||| |

Search completed: October 4, 2005, 22:31:06
Job time : 165 secs
```

THIS PAGE BLANK (USPTO)